



OIPE

#2

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/041,016

DATE: 02/16/2002
 TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt
 Output Set: N:\CRF3\02152002\J041016.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Jacobs, Kenneth
 6 McCoy, John M.
 7 Racie, Lisa A.
 8 LaVallie, Edward R.
 9 Merberg, David
 10 Treacy, Maurice
 11 Evans, Cheryl
 12 Agostino, Michael
 13 Lu, Zhijian
 14 Honjo, Tasuku
 W--> 15 Tashiro, Kei
 W--> 16 Nakamura, Tomoyuki
 18 (ii) TITLE OF INVENTION: SECRETED PROTEINS
 20 (iii) NUMBER OF SEQUENCES: 2
 22 (iv) CORRESPONDENCE ADDRESS:
 23 (A) ADDRESSEE: Genetics Institute, Inc.
 24 (B) STREET: 87 CambridgePark Drive
 25 (C) CITY: Cambridge
 26 (D) STATE: MA
 27 (E) COUNTRY: U.S.A.
 28 (F) ZIP: 02140
 30 (v) COMPUTER READABLE FORM:
 31 (A) MEDIUM TYPE: Floppy disk
 32 (B) COMPUTER: IBM PC compatible
 33 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 34 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 36 (vi) CURRENT APPLICATION DATA:
 C--> 37 (A) APPLICATION NUMBER: US/10/041,016
 C--> 38 (B) FILING DATE: 07-Jan-2002
 44 (C) CLASSIFICATION:
 C--> 41 (vii) PRIOR APPLICATION DATA:
 42 (A) APPLICATION NUMBER: US/09/083,002
 43 (B) FILING DATE: 21-MAR-1998
 45 (viii) ATTORNEY/AGENT INFORMATION:
 46 (A) NAME: Sprunger, Suzanne A.
 47 (B) REGISTRATION NUMBER: P-41,323
 49 (ix) TELECOMMUNICATION INFORMATION:
 50 (A) TELEPHONE: (617) 498-8284
 51 (B) TELEFAX: (617) 876-5851
 54 (2) INFORMATION FOR SEQ ID NO: 1:
 56 (i) SEQUENCE CHARACTERISTICS:

ENTERED

RAW SEQUENCE LISTING

DATE: 02/16/2002

PATENT APPLICATION: US/10/041,016

TIME: 13:04:10

Input Set : N:\Crif3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

```

57      (A) LENGTH: 2362 base pairs
58      (B) TYPE: nucleic acid
59      (C) STRANDEDNESS: double
60      (D) TOPOLOGY: linear
62      (ii) MOLECULE TYPE: cDNA
67      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
69 TAGCTTGGCA CGAGGGGACC CCGGCGCTCT CCCC GTGTCC TCTCCACGAC TCGCTCGGCC      60
71 CCTCTGGAAT AAAACACCCG CGAGCCCCGA GGGCCCAGAG GAGGCCGACG TGCCCGAGCT      120
73 CCTCCGGGGG TCCCGCCCGC GAGCTTTCTT CTCGCCTTCG CATCTCCTCC TCGCGCGTCT      180
75 TGGACATGCC AGGAATAAAA AGGATACTCA CTGTTACCAT TCTGGCTCTC TGTCTTCCAA      240
77 GCCCTGGGAA TGCACAGGCA CAGTGCACGA ATGGCTTTGA CCTGGATCGC CAGTCAGGAC      300
79 AGTGTTTAGA TATTGATGAA TGCCGAACCA TCCCCGAGGC CTGCCGAGGA GACATGATGT      360
81 GTGTTAACCA AAATGGCGGG TATTTATGCA TTCCCCGAC AAACCCTGTG TATCGAGGGC      420
83 CCTACTCGAA CCCCTACTCG ACCCCCTACT CAGGTCCGTA CCCAGCAGCT GCCCCACCAC      480
85 TCTCAGCTCC AAACATATCC ACGATCTCCA GGCCTCTTAT ATGCCGCTTT GGATACCAGA      540
87 TGGATGAAAG CAACCAATGT GTGGATGTGG ACGAGTGTGC AACAGATTCC CACCAGTGCA      600
89 ACCCCACCCA GATCTGCATC AATACTGAAG GCGGGTACAC CTGCTCCTGC ACCGACGGAT      660
91 ATTGGCTTCT GGAAGGCCAG TGCTTAGACA TTGATGAATG TCGCTATGGT TACTGCCAGC      720
93 AGCTCTGTGC GAATGTTCTT GGATCCTATT CTTGTACATG CAACCCTGGT TTTACCCTCA      780
95 ATGAGGATGG AAGGTCTTGC CAAGATGTGA ACGAGTGTGC CACCGAGAAC CCCTGCGTGC      840
97 AAACCTGCGT CAACACCTAC GGCTCTTTCA TCTGCCGCTG TGACCCAGGA TATGAACTTG      900
99 AGGAAGATGG CGTTCATTGC AGTGATATGG ACGAGTGCAG CTTCTCTGAG TTCCTCTGCC      960
101 AACATGAGTG TGTGAACCAG CCCGGCACAT ACTTCTGTCT CTGCCCTCCA GGCTACATCC      1020
103 TGCTGGATGA CAACCGAAGC TGCCAAGACA TCAACGAATG TGAGCACAGG AACCACACGT      1080
105 GCAACCTGCA GCAGACGTGC TACAATTTAC AAGGGGGCTT CAAATGCATC GACCCCATCC      1140
107 GCTGTGAGGA GCCTTATCTG AGGATCAGTG ATAACCGCTG TATGTGTCCT GCTGAGAACC      1200
109 CTGGCTGCAG AGACGAGCCC TTTACCATCT TGTACCGGGA CATGGACGTG GTGT CAGGAC      1260
111 GCTCCGTTCC CGCTAGACAT TTCCAATGC AAGCCACGAC CCGCTACCCT GGGGCCATT      1320
113 ACATTTTCCA GATCAAATCT GGGAATGAGG GCAGAGAATT TTACATGCGG CAAACGGGCC      1380
115 CCATCAGTGC CACCCTGGTG ATGACACGCC CCATCAAAGG GCCCCGGGAA ATCCAGCTGG      1440
117 ACTTGGAAT GATCACTGTC AACACTGTCA TCAACTTCAG AGGCAGCTCC GTGATCCGAC      1500
119 TGCGGATATA TGTGTGCGAG TACCCATTCT GAGCCTCGGG CTGGAGCCTC CGACGCTGCC      1560
121 TCTCATTGGC ACCAAGGGAC AGGAGAAGAG AGGAAATAAC AGAGAGAATG AGAGCGACAC      1620
123 AGACGTTAGG CATTTCTTGC TGAACGTTTC CCCGAAGAGT CAGCCCCGAC TTCCTGACTC      1680
125 TCACCTGTAC TATTGCAGAC CTGTCACCCT GCAGGACTTG CCACCCCCAG TTCCTATGAT      1740
127 ACAGTTATCA AAAAGTATTA TCATTGCTCC CCTGATAGAA GATTGTTGGT GAATTTTCAA      1800
129 GGCTTCAGT TTATTTCCAC TATTTTCAA GAAAATAGAT TAGGTTTGCG GGGGTCTGAG      1860
131 TCTATGTTCA AAGACTGTGA ACAGCTTGCT GTCACCTCTT CACCTCTTCC ACTCCTTCTC      1920
133 TCACTGTGTT ACTGCTTTGC AAAGACCCGG GAGCTGGCGG GGAACCCTGG GAGTAGCTAG      1980
135 TTTGCTTTTT GCGTACACAG AGAAGGCTAT GTAAACAAAC CACAGCAGGA TCGAAGGGTT      2040
137 TTTAGAGAA GTGTTTCAA ACCATGCCTG GTATTTTCAA CCATAAAAGA AGTTTCAGTT      2100
139 GTCCTTAAAT TTGTATAACG GTTTAATTCT GTCTTGTTCA TTTTGAGTAT TTTTAAAAA      2160
141 TATGTCGTAG AATTCCTTCG AAAGGCCTTC AGACACATGC TATGTTCTGT CTTCCCAAAC      2220
143 CCAGTCTCCT CTCCATTTA GCCCAGTGTT TTCTTTGAGG ACCCCTTAAT CTTGCTTTCT      2280
145 TTAGAATTTT TACCAATTG GATTGGAATG CAGAGGTCTC CAAACTGATT AAATATTTGA      2340
147 AGAGAAAAAA AAAAAAAAAA AA      2362
149 (2) INFORMATION FOR SEQ ID NO: 2:
151      (i) SEQUENCE CHARACTERISTICS:
152      (A) LENGTH: 448 amino acids

```

RAW SEQUENCE LISTING

DATE: 02/16/2002

PATENT APPLICATION: US/10/041,016

TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

```

153      (B) TYPE: amino acid
154      (C) STRANDEDNESS: Not Relevant
155      (D) TOPOLOGY: linear
157      (ii) MOLECULE TYPE: protein
162      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
164      Met Pro Gly Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys
165      1          5          10          15
167      Leu Pro Ser Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp
168      20          25          30
170      Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr
171      35          40          45
173      Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly
174      50          55          60
176      Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr
177      65          70          75          80
179      Ser Asn Pro Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala
180      85          90          95
182      Pro Pro Leu Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile
183      100         105         110
185      Cys Arg Phe Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val
186      115         120         125
188      Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys
189      130         135         140
191      Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp
192      145         150         155         160
194      Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr
195      165         170         175
197      Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys
198      180         185         190
200      Asn Pro Gly Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val
201      195         200         205
203      Asn Glu Cys Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr
204      210         215         220
206      Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu
207      225         230         235         240
209      Asp Gly Val His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe
210      245         250         255
212      Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser
213      260         265         270
215      Cys Pro Pro Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp
216      275         280         285
218      Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr
219      290         295         300
221      Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys
222      305         310         315         320
224      Glu Glu Pro Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala
225      325         330         335
227      Glu Asn Pro Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp
228      340         345         350

```

RAW SEQUENCE LISTING

DATE: 02/16/2002

PATENT APPLICATION: US/10/041,016

TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

230	Met	Asp	Val	Val	Ser	Gly	Arg	Ser	Val	Pro	Ala	Asp	Ile	Phe	Gln	Met
231			355						360				365			
233	Gln	Ala	Thr	Thr	Arg	Tyr	Pro	Gly	Ala	Tyr	Tyr	Ile	Phe	Gln	Ile	Lys
234			370				375					380				
236	Ser	Gly	Asn	Glu	Gly	Arg	Glu	Phe	Tyr	Met	Arg	Gln	Thr	Gly	Pro	Ile
237			385			390					395				400	
239	Ser	Ala	Thr	Leu	Val	Met	Thr	Arg	Pro	Ile	Lys	Gly	Pro	Arg	Glu	Ile
240				405						410					415	
242	Gln	Leu	Asp	Leu	Glu	Met	Ile	Thr	Val	Asn	Thr	Val	Ile	Asn	Phe	Arg
243				420					425					430		
245	Gly	Ser	Ser	Val	Ile	Arg	Leu	Arg	Ile	Tyr	Val	Ser	Gln	Tyr	Pro	Phe
246			435				440						445			

VERIFICATION SUMMARY

DATE: 02/16/2002

PATENT APPLICATION: US/10/041,016

TIME: 13:04:11

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

L:37 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:38 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:41 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:15 M:259 W: Allowed number of lines exceeded, (i) APPLICANT:
L:16 M:259 W: Allowed number of lines exceeded, (i) APPLICANT: